



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
| 09/817,591      | 03/26/2001  | Yihong Gong          | CA1122              | 7751             |

23493 7590 04/20/2005

SUGHRUE MION, PLLC  
401 Castro Street, Ste 220  
Mountain View, CA 94041-2007

|          |
|----------|
| EXAMINER |
|----------|

TRAN, QUOC A

|          |              |
|----------|--------------|
| ART UNIT | PAPER NUMBER |
|----------|--------------|

2176

DATE MAILED: 04/20/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

# Office Action Summary

Application No.

09/817,591

Applicant(s)

GONG ET AL.

Examiner

Quoc A. Tran

Art Unit

2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☒ Responsive to communication(s) filed on 22 December 2004.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-32 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-32 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- |   |   |
|---|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)  | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date <u>02/12/2002</u> . | 6) <input type="checkbox"/> Other: _____  |

### DETAILED ACTION

1. This action is responsive to Amendment A filed 12/22/2004.
2. Claims 1-32 are pending. Claims 1, 9, 13, 21, 26 and 29 are independent claims.

### ***Claim Rejections - 35 USC § 102***

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. **Claims 1, 9 are rejected under 35 U.S.C. 102(e) as being anticipated by Billheimer et al. US Patent No. 6,611,825 B1 - filed 06/09/1999 (hereinafter '825).**

**In regard to independent claim 1, “creating a weighted document term-frequency vector for said document ...”,** as taught by '825 at col. 6, lines 10-25 (i.e.... information retrieval can be performed upon user request. Information retrieval is performed by constructing a term frequency query vector ... determining similarity by measuring the distance between the query and the documents, and returning ranked matching documents ...),

*"for each sentence in said document, creating a weighted sentence term-frequency vector"*, as taught by '825 at col. 6, lines 20-35 (i.e.... The term frequency vector is constructed by tokenizing using the same policy as was used for the original document collection, performing any optional functions using the same policy as was used for the original document collection, and identifying the terms in the original term frequency matrix representing the original document collection ..... document cross-referencing can be performed by segmenting the documents of the document collection into a plurality of suitably sized units, constructing a subspace representation for the units from all or a subset of the documents, projecting all the units into that subspace, determining similarity by measuring the distance between the projections of the units from one or more documents and the projections of the units in one or more target documents, and returning ranked matching units as potential cross-reference sections ... ). Examiner reads segmenting the documents of the document collection into a plurality of suitably sized units, and determining similarity by measuring, which could be interpreted as the claimed *"each sentence in said document, creating a weighted..."*.

*"computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector"*, as taught by '825 at col. 17, lines 25-35 (i.e.... FIG. 15 is a pictorial representation of a term frequency matrix ... to the right of the table is the query vector. At the bottom of each column is the score for the query vector...),

*"selecting a sentence for inclusion in said generic text summary in accordance with said computing"*, as taught by '825 at col. 5, lines 45-55 (i.e.... document

summarization can be performed for an individual document, i.e., a few words or terms summarizing the contents of a document or indicating its topical affiliations, can be provided by: projecting the document into a subspace; projecting the document projection back into a term space; identifying a plurality of terms with the largest entries; and returning the identified terms as the document summarization...).

**In regard to independent claim 9**, is directed to a computer system for performing the method of claim 1 and is similarly rejected along the same rationale.

***Claim Rejections - 35 USC § 103***

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. **Claims 2-8, 10-32 are rejected under 35 U.S.C. 103(a) as being unpatentable by Billheimer et al. US Patent No. 6,611,825 B1 - filed 06/09/1999 (hereinafter '825), in view of Herz US Patent No. 6,029,195 - filed 12/05/1997 (hereinafter '195).**

**In regard to dependent claim 2**, "*creating a weighted document term-frequency vector for said document ...*", as taught by '825 at col. 6, lines 10-25 (i.e.... information retrieval can be performed upon user request. Information retrieval is performed by

constructing a term frequency query vector ... determining similarity by measuring the distance between the query and the documents, and returning ranked matching documents...).

'825 does not explicitly teach, "*recreating said weighted document term-frequency vector in accordance with said deleting and said eliminating; and selectively repeating said computing, said selecting, said deleting, said eliminating, and said recreating*", however, as taught by '195 at col. 13, lines 40-67 (i.e.... textual and associative attributes ... for information retrieval purposes they can be decomposed into smaller, simpler numeric attributes... In particular, a textual attribute, such as the full text of a movie review, can be replaced by a collection of numeric attributes that represent scores to denote the presence and significance of the words ... score is often called the "term frequency" (TF) of the word...occurrence of a word in the text's title might be counted as a 3-fold or more generally k-fold occurrence...). Examiner reads: replaced by a collection of ..., and ...k-fold occurrence ..., which could be interpreted as the claimed "*recreating...*", and "*repeating...*".

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein recreating weighted document term-frequency vector in accordance with deleting and eliminating; and selectively repeating the process. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication

Art Unit: 2176

network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 3, "*selectively repeating is terminated when a predetermined number of sentences has been selected*",** as taught by '195 at col. 13, line 40 through col. 14, line 35 (i.e.... textual and associative attributes ... for information retrieval purposes they can be decomposed into smaller, simpler numeric attributes... In particular, a textual attribute, such as the full text of a movie review, can be replaced by a collection of numeric attributes that represent scores to denote the presence and significance of the words ... score is often called the "term frequency" (TF) of the word... occurrence of a word in the text's title might be counted as a 3-fold or more generally k-fold occurrence... break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams. ... Conceptually speaking, in the character 5-gram case, the textual attribute would be decomposed into at least  $26^{sup.5} = 11,881,376$  numeric attributes), also as taught by '195 at col. 70, lines 40-60 (i.e.... The user first chooses one of the highest level (largest) clusters from a menu, and is presented with a menu listing the subclusters of said cluster ... This process is repeated until the user comes to a leaf of the tree... ).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein selectively repeating is terminated when a predetermined number of sentences has been selected. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e...The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 4**, *"calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector"*, however, as taught by '195 at col. 15, line 45 through col. 16, line 45 (i.e.... value  $V$  may be decomposed as described above into a collection of real numbers, representing the association scores between the target object in question and various ancillary objects.  $V$  may therefore be regarded as a vector with components  $V_{sub.1}$ ,  $V_{sub.2}$ ,  $V_{sub.3}$ , etc., representing the association scores between the object and ancillary objects 1, 2, 3, etc., respectively. The distance between two vector values  $V$  and  $U$  of an associative attribute is then computed using the angle distance measure,  $\arccos(V \cdot U / \sqrt{(V \cdot V)(U \cdot U)})$ . (Note that the three inner products in this



expression have the form  $XY_{sup.t} = X_{sub.1} Y_{sub.1} + X_{sub.2} Y_{sub.2} + X_{sub.3} Y_{sub.3} + \dots$ , and that for efficient computation, terms of the form  $X_{sub.i} Y_{sub.i}$  may be omitted from this sum if either of the scores  $X_{sub.i}$  and  $Y_{sub.i}$  is zero.) ...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 5**, *"creating a weighted sentence term-frequency vector comprises implementing a local weighting function and implementing a global weighting function"*, as taught by '195 at col. 60, lines 35-55 (i.e.... global request messages, step 4 eventually causes all core servers for topic  $C_k$  to act on request  $R$  and therefore store a local copy of file  $F$ ... it also computes and stores the weight  $w(S_i,$

C.sub.F), where C.sub.F is a cluster consisting of the single target object associated with file F ...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein creating a weighted sentence term-frequency vector comprises implementing a local weighting function and implementing a global weighting function. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing...).

**In regard to dependent claim 6,** *"creating a weighted sentence term-frequency vector comprises normalizing each said weighted sentence term-frequency vector"*, as taught by '825 at col. 5, lines 10-250 (i.e.... when the document collection is a dynamically changing document collection, the subspace is updated as the document collection changes...new documents, constructing a new term frequency matrix, statistically transforming the term frequencies in the same way that the initial document collection was transformed, projecting the term frequency matrix on the term subspace,

Art Unit: 2176

computing the residual, augmenting the term subspace with the normalized residual, expanding the document subspace, and re-identifying significant features in the subspace...).

**In regard to dependent claim 7**, *"creating a weighted document term-frequency vector comprises implementing a local weighting function and implementing a global weighting function"*, as taught by '195 at col. 60, lines 35-55 (i.e.... global request messages, step 4 eventually causes all core servers for topic C<sub>k</sub> to act on request R and therefore store a local copy of file F... it also computes and stores the weight  $w(S_i, C_{sub.F})$ , where C<sub>sub.F</sub> is a cluster consisting of the single target object associated with file F ...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein creating a weighted document term-frequency vector comprises implementing a local weighting function and implementing a global weighting function. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity

of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing...).

**In regard to dependent claim 8**, "*creating a weighted document term-frequency vector comprises normalizing each said weighted document term-frequency vector*", as taught by '825 at col. 5, lines 10-250 (i.e.... when the document collection is a dynamically changing document collection, the subspace is updated as the document collection changes...new documents, constructing a new term frequency matrix, statistically transforming the term frequencies in the same way that the initial document collection was transformed, projecting the term frequency matrix on the term subspace, computing the residual, augmenting the term subspace with the normalized residual, expanding the document subspace, and re-identifying significant features in the subspace...).

**In regard to dependent claims 10**, is directed to a computer system for performing the method of claim 2 and is similarly rejected along the same rationale.

**In regard to dependent claim 11**, is directed to a computer system for performing the method of claims 2-3 and is similarly rejected along the same rationale.

**In regard to dependent claim 12**, is directed to a computer system for performing the method of claim 3 and is similarly rejected along the same rationale.

**In regard to independent claim 13**, incorporate substantially similar subject matter as cited in claims 1-3 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 14**, incorporate substantially similar subject matter as cited in claim 2 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 15**, incorporate substantially similar subject matter as cited in claim 3 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 16**, incorporate substantially similar subject matter as cited in claim 4 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 17**, incorporate substantially similar subject matter as cited in claim 5 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 18**, incorporate substantially similar subject matter as cited in claim 6 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 19**, incorporate substantially similar subject matter as cited in claim 7 above, and is similarly rejected along the same rationale.

**In regard to dependent claim 20**, incorporate substantially similar subject matter as cited in claim 8 above, and is similarly rejected along the same rationale.

**In regard to independent claim 21**, "*selecting a sentence for inclusion in said generic text summary in accordance with said ranking*", as taught by '825 at col. 5, lines 45-55 (i.e.... document summarization can be performed for an individual document, i.e., a few words or terms summarizing the contents of a document or indicating its topical affiliations, can be provided by: projecting the document into a subspace; projecting the document projection back into a term space; identifying a plurality of terms with the largest entries; and returning the identified terms as the document summarization...).

'825 does not explicitly teach, "*constructing a terms-by-sentences matrix for said document; performing singular value decomposition on said terms-by-sentences matrix*

*to obtain a singular value matrix and a right singular vector matrix, wherein each sentence in said document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix*", however, as taught by '195 at col. 16, lines 40-65 (i.e.... the matrix A is the dimensionality-reducing linear transformation (or an approximation thereto) determined by collecting the vector values of the textual attribute, for all target objects known to the system, and applying singular value decomposition to the resulting collection. The same approach can be applied to the vector values of associative attributes. The above definitions allow us to determine how close together two target objects are with respect to a single attribute, whether numeric, associative, or textual. The distance between two target objects X and Y with respect to their entire multi-attribute profiles P.sub.X and P.sub.Y is then denoted d(X,Y) or d(P.sub.X, P.sub.Y) and defined as: ... (((distance with respect to attribute a)(weight of attribute a)).sup.k +((distance with respect to attribute b)(weight of attribute b)).sup.k +((distance with respect to attribute c)(weight of attribute c)).sup.k + . . . ).sup.k...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein constructing a terms-by-sentences matrix for a document; performing singular value decomposition on said terms-by-sentences matrix to obtain a singular value matrix and a right singular vector matrix, wherein each sentence in the document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix. One of ordinary

Art Unit: 2176

skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 22, "*repeating said selecting*",** as taught by '195 at col. 70, lines 40-60 (i.e.... The user first chooses one of the highest level (largest) clusters from a menu, and is presented with a menu listing the subclusters of said cluster ... This process is repeated until the user comes to a leaf of the tree... ).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein creating a generic text summary of a document; said method comprising repeating said selecting. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to

the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 23, "*repeating is terminated when a predetermined number of sentences has been selected*",** as taught by '195 at col. 13, line 40 through col. 14, line 35 (i.e.... textual and associative attributes ... for information retrieval purposes they can be decomposed into smaller, simpler numeric attributes... In particular, a textual attribute, such as the full text of a movie review, can be replaced by a collection of numeric attributes that represent scores to denote the presence and significance of the words ... score is often called the "term frequency" (TF) of the word... occurrence of a word in the text's title might be counted as a 3-fold or more generally k-fold occurrence... break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams. ... Conceptually speaking, in the character 5-gram case, the textual attribute would be decomposed into at least  $26^{sup.5} = 11,881,376$  numeric attributes) also as taught by '195 at col. 70, lines 40-60 (i.e.... The user first chooses one of the highest level (largest) clusters from a menu, and is presented with a menu listing the subclusters of said cluster ... This process is repeated until the user comes to a leaf of the tree... ).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein selectively repeating is terminated when a predetermined number of sentences has been selected. One of ordinary skill in the art would have been motivated



Art Unit: 2176

to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 24**, *"wherein said selecting further comprises identifying a sentence having a desired index value with said right singular vector"*, however, as taught by '195 at col. 15, line 45 through col. 16, line 45 (i.e.... value  $V$  may be decomposed as described above into a collection of real numbers, representing the association scores between the target object in question and various ancillary objects.  $V$  may therefore be regarded as a vector with components  $V_{\text{sub.1}}$ ,  $V_{\text{sub.2}}$ ,  $V_{\text{sub.3}}$ , etc., representing the association scores between the object and ancillary objects 1, 2, 3, etc., respectively. The distance between two vector values  $V$  and  $U$  of an associative attribute is then computed using the angle distance measure,  $\arccos(VU_{\text{sup.t}} / \sqrt{(Vv_{\text{sup.t}})(UU_{\text{sup.t}})})$ . (Note that the three inner products in this expression have the form  $XY_{\text{sup.t}} = X_{\text{sub.1}} Y_{\text{sub.1}} + X_{\text{sub.2}} Y_{\text{sub.2}} + X_{\text{sub.3}} Y_{\text{sub.3}} + \dots$ , and that for efficient computation, terms of the form  $X_{\text{sub.i}} Y_{\text{sub.i}}$  may be omitted from this sum if either of the scores  $X_{\text{sub.i}}$  and  $Y_{\text{sub.i}}$  is zero.) ... ).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing... ).

**In regard to dependent claim 25,** *"wherein said constructing comprises implementing a local weighting function and implementing a global weighting function"*, as taught by '195 at col. 60, lines 35-55 (i.e.... global request messages, step 4 eventually causes all core servers for topic C<sub>k</sub> to act on request R and therefore store a local copy of file F... it also computes and stores the weight  $w(S_i, C_{\text{sub}}.F)$ , where C<sub>sub</sub>.F is a cluster consisting of the single target object associated with file F ...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified '195 into '825 to provide a way to include the feature, wherein creating a weighted sentence term-frequency vector comprises

implementing a local weighting function and implementing a global weighting function.

One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing...).

**In regard to independent claim 26**, is directed to a computer system for performing the method of claim 1, 21 and is similarly rejected along the same rationale.

**In regard to dependent claims 27-28 consecutively**, are directed to a computer system for performing the method of claims 22-23 consecutively and are similarly rejected along the same rationale.

**In regard to independent claim 29**, incorporate substantially similar subject matter as cited in claims 13, 21 above, and is similarly rejected along the same rationale.

**In regard to dependent claims 30-32 consecutively**, incorporate substantially similar subject matter as cited in claims 23-25 consecutively, and are similarly rejected along the same rationale.

***Response to Argument***

7. Examiner has completed a through study of Applicant's Remark filed 12/22/2004; Applicant's arguments have been fully considered but they are not persuasive. The reason for rejection was set forth in the previous rejection, a copy is cited above.

**Reponses to Remarks pages 1-3:**

In regard to the rejection of claim 1 and 9, Applicant argues that, Billheimer failed to teach or suggest the features in claim 1 of the invention. The Office respectfully disagreed, for more detail sees the previous rejection, which set forth in the rejection above.

Further more Applicant argues that, Billheimer failed to teach the feature, wherein creating a generic text summary of a document. The Office respectfully disagreed, for more detail sees the previous rejection, which set forth in the rejection above, and also as taught by Billheimer at the Abstract (i.e. A text mining program is provided that allows a user to perform text mining operations, such as: information retrieval, term and document visualization, term and document clustering, term and document classification, summarization of individual documents and groups of documents, and document cross-referencing. This is accomplished by representing the text of a document collection using subspace transformations. This subspace transformation representation is performed by: constructing a term frequency matrix of the term frequencies for each of the documents, transforming the term frequencies for statistical purposes, and projecting the documents or the terms into a lower dimensional

subspace. As the document collection is updated, the subspace is dynamically updated to reflect the new document collection).

Further more Applicant argues that, Billheimer failed to teach the feature, computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector. The Office respectfully disagreed, for more detail sees the previous rejection, which set forth in the rejection above, and also as taught by Billheimer at col. 11, line 20 through col. 18, line 35, wherein Billheimer disclosed in detail the process of computational of equation (1), (2), (3), (7) and (8) resulting the decomposition result of a term basis matrix, a weight matrix and a document basis matrix. Also disclosed in col. 22, lines 25-45, wherein constructing of the term frequency query vector such as identifying the terms that are in the original term frequency matrix representing the original document collection; and determining a similarity measurement between the text unit and the subspace representations of the text units in one or more target documents for returning a plurality of ranked matching text units based on the determined similarity measurement of term.

There for independent claim 1 remain rejected.

**Reponses to Remarks page 4:**

In regard to the rejection of claim 2-8, Applicant argues that, Herz failed to teach or suggest the weighted-sentence term-frequency, and failed to provide the deficiencies of Billheimer. The Office respectfully disagreed, for more detail sees the previous rejection, which set forth in the rejection above, and also noticed that, the Office

Art Unit: 2176

recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. In this case,

Herz did not explicitly teaches, weighted-sentence term-frequency; however as taught by Billheimer at col. 6, lines 10-25 (i.e.... information retrieval can be performed upon user request. Information retrieval is performed by constructing a term frequency query vector ... determining similarity by measuring the distance between the query and the documents, and returning ranked matching documents...).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Herz' s teaching into Billheimer's teaching to provide a way to include the feature, wherein calculating an inner product of said weighted sentence term-frequency vector and said weighted document term-frequency vector. One of ordinary skill in the art would have been motivated to perform such a modification to provides users the ability for filtering, browsing and for pre-caching of data at locations on the data communication network and at times that minimize the traffic flow in the communication network to thereby efficiently provide the desired information to the user and/or conserve valuable storage space by only storing those target objects (or segments thereof) which are relevant to the user's interests, as taught by '195 at col. 7, line 50 through col. 8, line 50 ( i.e... The ability to measure the similarity of profiles describing target objects and a user's interests can be applied in two basic ways: filtering and browsing...).

There for dependent claims 2-8 remain rejected.

**Reponses to Remarks page 5:**

In regard to the rejection of claim 9-12, is directed to a computer system for performing the method of claims 1-3 and are similarly rejected along the same rationale.

**Reponses to Remarks page 5:**

In regard to the rejection of claim 13, incorporate substantially similar subject matter as cited in claims 1-3 above, and is similarly rejected along the same rationale.

**Reponses to Remarks pages 5-7:**

In regard to the rejection of claims 21-25, 26-28 and 29-32, Applicant argues that, non of the cite preferences teach Billheimer failed to teach the feature,

terms-by-sentences matrix for said document. The Office respectfully disagreed, for more detail sees the previous rejection, which set forth in the rejection above, and also as taught by Herz at col. 13, line 40 through col. 14, line 35 (i.e.... textual and associative attributes ... for information retrieval purposes they can be decomposed into smaller, simpler numeric attributes... In particular, a textual attribute, such as the full text of a movie review, can be replaced by a collection of numeric attributes that represent scores to denote the presence and significance of the words ... score is often called the "term frequency" (TF) of the word...occurrence of a word in the text's title might be counted as a 3-fold or more generally k-fold occurrence... break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams. ... Conceptually speaking, in the character 5-gram case, the textual attribute would be decomposed into at least  $26^{sup.5} = 11,881,376$  numeric attributes), also as

taught by Billheimer at the Abstract (i.e. A text mining program is provided that allows a user to perform text mining operations, such as: information retrieval, term and document visualization, term and document clustering, term and document classification, summarization of individual documents and groups of documents, and document cross-referencing. This is accomplished by representing the text of a document collection using subspace transformations. This subspace transformation representation is performed by: constructing a term frequency matrix of the term frequencies for each of the documents, transforming the term frequencies for statistical purposes, and projecting the documents or the terms into a lower dimensional subspace. As the document collection is updated, the subspace is dynamically updated to reflect the new document collection).

There for dependent claims 21-25, 26-28 and 29-32 remain rejected.

### ***Conclusion***

8. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of



Art Unit: 2176

the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Quoc A. Tran whose telephone number is (571) 272-4103. The examiner can normally be reached on Monday through Friday from 8:30AM to 5:00PM.

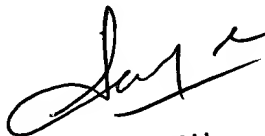
If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joseph H. Feild can be reached on (571) 272-4090. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

---

Quoc A. Tran  
Patent Examiner  
Technology Center 2176  
April 13, 2005

  
SANJIV SHAH  
PRIMARY EXAMINER